The Bounded-Storage Model in the Presence of a Quantum Adversary

Robert T. König and Barbara M. Terhal

***Abstract*—An extractor is a function E that is used to extract randomness. Given an imperfect random source X and a uniform seed Y, the output $\mathsf{E}(X,Y)$ is close to uniform. We study properties of such functions in the presence of prior quantum information about X, with a particular focus on cryptographic applications. We prove that certain extractors are suitable for key expansion in the bounded-storage model where the adversary has a limited amount of quantum memory. For extractors with one-bit output we show that the extracted bit is essentially equally secure as in the case where the adversary has classical resources. We prove the security of certain constructions that output multiple bits in the bounded-storage model.**

***Index Terms*—Bounded-storage model, cryptography, extractors, locking, privacy amplification, quantum information theory, quantum key distribution, quantum memory, security proofs, universal composability.**

I. Introduction

THE aim of *randomness extraction* is to generate "almost uniform" randomness given an imperfect source of randomness X. The term "extractor" is generally used to describe a procedure which accomplishes this task; more formally, an extractor is a (deterministic) function $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ which, when applied to an imperfect source X and a uniform and independent seed Y, yields an output $Z := \mathsf{E}(X,Y)$ which is close to being uniformly distributed on $\mathcal{Z}$. Such an extractor is characterized by a number of parameters. Among these are the amount of randomness Y that is required, the amount of randomness Z produced, and, most importantly, the character of the sources X which lead to almost uniform output. A very general class of sources are the *weak* sources X, characterized by a lower bound on the *min-entropy* $H_\infty(X) := -\log \max_x P_X(x)$. Correspondingly, a (k,ε)-extractor [1] commonly refers to an extractor which, for any input distribution P_X with $H_\infty(X) \geq k$, outputs ε-uniform randomness Z.

Besides purifying randomness, extractors are an essential tool in computer science, in particular in complexity theory and cryptography. Correspondingly, the study of such extractors has been a major research topic in recent years, and much understanding has been gained (see [2] for a review). For applications in computer science, the challenge is to find explicit, efficiently computable extractors with good parameters.

In a cryptographic context, a certain variant of the concept of a (k,ε)-extractor is of particular importance. These are called *strong* extractors; they have the additional property that even the pair $(Y,\mathsf{E}(X,Y))$ is ε-close to uniform. This means for example that $(Y,\mathsf{E}(X,Y))$ can be used to encrypt a message $M = (M_1, M_2)$ using a one-time pad [3] as $C = (C_1, C_2) = (M_1 \oplus Y, M_2 \oplus \mathsf{E}(X,Y))$. An adversary who learns the cipher-text C as well as the message M_1 (and thus the seed Y) will be completely ignorant of the content of the remaining message M_2. Expressed differently, the pair $(Y,\mathsf{E}(X,Y))$ is a key with *universally composable* security [4], [5].

A more striking application of strong extractors in cryptography is privacy amplification, introduced by Bennett, Brassard, and Robert [6] and further analyzed in [7]. This refers to a technique that allows two parties, Alice and Bob, to generate a secret key Z from a shared random variable X about which the adversary has partial information E. The only assumption is that the parties are connected by an authentic but otherwise completely insecure channel. The key Z is then obtained as follows: Alice generates an independent uniform seed Y and sends it over the channel. Subsequently, both parties apply a strong extractor to get $Z := \mathsf{E}(X,Y)$. The security of Z when used as a secret key directly follows from the properties of the strong extractor, assuming a certain bound on the information E of the adversary.

Apparently related to privacy amplification, but conceptually quite different, is Maurer's bounded-storage model [8]. The first security proof for general adversaries in this model was obtained by Aumann, Ding, and Rabin [9] and essentially optimal constructions were subsequently found in a sequence of papers [10]–[12]. The aim of the honest parties in the bounded-storage model is not *key extraction* but *key expansion*. In this setting, a large amount of randomness X is publicly, but only temporarily available. Alice and Bob use a previously shared (short) secret key Y to obtain additional key bits $Z = \mathsf{E}(X,Y)$ using a strong extractor. The seed Y remains hidden to the adversary until (possibly) after the execution of the protocol. The adversary is assumed to have only a bounded amount of storage (which may be much larger than the honest parties' memory). As a result, his information E about X is limited, once X becomes inaccessible, and by the properties of the extractor, Z can be shown to be secure even if he later obtains the seed Y (this was referred to as "everlasting security" in [9]).

From a cryptographic viewpoint, a natural generalization of these scenarios is arrived at by allowing the adversary to have

Manuscript received September 13, 2006; revised March 21, 2007. The work of R. T. König was supported by the European Commission through the FP6-FET Integrated Project SCALA, CT-015714. The work of B. M. Terhal was supported by the NSA and the ARDA through ARO Contract W911NF-04-C-0098. The material in this paper was presented at QIP 2007, Brisbane, Australia, January 2007.

R. T. König is with the Institute for Quantum Information, California Institute of Technology, Pasadena, CA 91125 USA.

B. M. Terhal is with the IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598 USA.

Communicated by A. Winter, Associate Editor for Quantum Information Theory.

Digital Object Identifier 10.1109/TIT.2007.913245

quantum information Q instead of only classical information E about X. This modification is not merely of theoretical interest. Indeed, the only construction proved to be secure [13]–[16] for privacy amplification has found various applications in quantum cryptography. Besides simplifying and improving security proofs for quantum key distribution [16], [17], the quantum version of privacy amplification has been used to derive both possibility [18] and impossibility [19] results for tasks such as bit commitment or oblivious transfer.

While the problem of constructing strong extractors is well-studied, little is known about the security resulting from their use in a quantum context. For the bounded-storage model, Gavinsky, Kempe, and de Wolf [20] recently gave an example of an extractor which yields a classically secure key, but is completely insecure against an adversary with a similar amount of quantum storage. There is no construction for the bounded-storage model that is known to be secure against a quantum adversary.

In this paper, we study properties of strong extractors in a context where the adversary has quantum information. Our main focus is on the two cryptographic settings described. We give the first constructions of extractors that are usable in the bounded-storage model against a quantum adversary, and we show that certain strong extractors generate secure key bits in the setting of privacy amplification. This reduces the amount of randomness needed in certain applications.

Outline: In Section II, we introduce the relevant definitions. In Section III, we show that any strong extractor which outputs a single bit yields essentially the same degree of security in a cryptographic setting, irrespective of whether the adversary has quantum or classical information. We then use a hybrid argument in Section IV to obtain extractors that output several bits. In Section V, we explain how these extractors can be used in the bounded-storage model. Finally, we show that general strong extractors can be used in the setting of privacy amplification in Section VI. We conclude in Section VII.

A. Notation

Throughout this paper, all logarithms are binary, i.e., to base 2. For a random variable X with range $\mathcal{X}$, we define the *min-entropy* of X as $H_\infty(X) := -\log \max_x P_X(x)$. More generally, for a quantum state ρ_Q on a Hilbert space $\mathcal{Q}$, $H_\infty(Q)$ is the min-entropy of the distribution of eigenvalues of ρ_Q. Analogously, the max-entropy is defined as $H_0(X) := \log|\mathrm{supp}(P_X)| = \log|\mathcal{X}|$ and $H_0(Q) := \log \mathrm{rank}(\rho_Q)$, respectively. Expressed differently, $H_0(Q)$ can be understood as the number of qubits constituting system Q. For a function $g : \mathcal{X} \to \mathbb{R}$, we denote by

$$\mathbb{E}_{x \leftarrow P_X}[g(x)] := \sum_{x \in \mathcal{X}} P_X(x) g(x)$$

the expectation of $g(X)$ over a random choice of $x \leftarrow P_X$. We also use the notation $P_X \cdot P_Y$ to refer to the joint distribution of two independent random variables X and Y, that is, $\Pr[X = x, Y = y] = P_X(x) \cdot P_Y(y)$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

In the sequel, Q refers to a quantum system, whereas E, V, W, X, Y, and Z are assumed to be classical. Slightly abusing notation, we sometimes refer to the Hilbert space corresponding to a classical-quantum state (cq-state) ρ_{XQ} by $\mathcal{X} \otimes \mathcal{Q}$. We denote the completely mixed state on $\mathcal{X}$ by $\rho_{\mathcal{U}_\mathcal{X}}$.

We will sometimes use cq-states with multipartite classical parts, e.g., a ccq-state ρ_{XYQ}. For such a state ρ_{XYQ}, we say that $Y \leftrightarrow X \leftrightarrow Q$ forms a Markov chain if it has the form

$$\rho_{XYQ} = \sum_{x,y} P_{XY}(x,y) |xy\rangle\langle xy| \otimes \rho_x \tag{1}$$

for some states $\{\rho_x\}_{x \in \mathcal{X}}$ on $\mathcal{Q}$. A state with this property defines a distribution P_{XY}, which defines the conditional distributions $P_{X\,|\,Y=y}$ and, for any function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, the distribution $P_{f(X,Y)XY}$. The corresponding conditional states $\rho_{XQ\,|\,Y=y}$ are obtained by making the appropriate replacement in (1), i.e.,

$$\rho_{XQ\,|\,Y=y} = \sum_x P_{X\,|\,Y=y}(x) |x\rangle\langle x| \otimes \rho_x.$$

Similarly, we can define the cccq-state

$$\rho_{f(X,Y)XYQ} = \sum_{x,y} P_{XY}(x,y) |f(x,y)xy\rangle\langle f(x,y)xy| \otimes \rho_x$$

which in turn gives rise to states such as $\rho_{f(X,y)XQ\,|\,Y=y}$.

We will use the trace norm $||A|| := \frac{1}{2}\mathrm{tr}(\sqrt{A^\dagger A})$ for any operator A. We include the factor $\frac{1}{2}$ in this definition for convenience. It ensures that the distance $||\rho - \sigma||$ of two states ρ and σ is in the interval $[0, 1]$. Note that if ρ_{XQ} and $\sigma_{X'Q'}$ are cq-states on $\mathcal{X} \otimes \mathcal{Q}$, then

$$||\rho_{XQ} - \sigma_{X'Q'}|| = \sum_{x \in \mathcal{X}} ||P_X(x)\rho_x - P_{X'}(x)\sigma_x||. \tag{2}$$

For two probability distributions P and Q on $\mathcal{X}$, the trace norm of their difference (when identifying the distribution with a state via an orthonormal basis), i.e.,

$$||P - Q|| := \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)|$$

is also known as the variational distance.

Let $\rho_{XQ} = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x| \otimes \rho_x$ be a cq-state. Consider a fixed positive operator-valued measure (POVM) $\mathcal{E} := \{E_z\}_{z \in \mathcal{Z}}$ on $\mathcal{Q}$. We denote by $P_{XZ} \equiv \rho_{X\mathcal{E}(Q)}$ the joint distribution of X and the measurement outcome, i.e.,

$$P_{Z\,|\,X=x}(z) = \mathrm{tr}(E_z \rho_x)$$

for every $z \in \mathcal{Z}$ and $x \in \mathcal{X}$.

We will often encounter scalar quantities d that are functions of a given distribution or a quantum state, i.e., $d = d(P_X)$ or $d = d(\rho_Q)$. In these cases, we use the shorthand $d(X)$ or $d(Q)$. Similarly, we write $d(Q\,|\,W = w)$ instead of $d(\rho_{Q\,|\,W=w})$. More generally, we will consider quantities that depend on a specific bipartition of a state ρ_{ZE} into Z and E; in these cases, we write $d(Z \leftarrow E)$. Again, we use the notation $d(Z \leftarrow E\,|\,W = w)$ to denote the corresponding quantity for the conditional state $\rho_{ZE\,|\,W=w}$.

II. Extractors and Secret Keys

A. Classical Adversaries

Before reviewing the definition of strong extractors and a number of their basic properties, let us introduce a shorthand notation for the *nonuniformity*, a quantity which measures the extent to which a probability distribution of a random variable Z deviates from the uniform distribution, possibly given another random variable E:

Definition 1: Let P_{ZE} be an arbitrary distribution. The *nonuniformity* $d(Z \leftarrow E)$ *of* Z *given* E is defined as

$$d(Z \leftarrow E) := \|P_{ZE} - P_{\mathcal{U}_\mathcal{Z}} \cdot P_E\|.$$

Here P_E is the marginal distribution of P_{ZE}, and $P_{\mathcal{U}_\mathcal{Z}}$ denotes the uniform distribution on $\mathcal{Z}$.

Note that $d(Z)$ is simply the distance of the distribution P_Z from the uniform distribution. A strong extractor can then be defined as follows.

Definition 2: A *strong* (k,ε)*-extractor* is a function $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ with the property that

$$d(\mathsf{E}(X,Y) \leftarrow Y) = \left\|P_{\mathsf{E}(X,Y)Y} - P_{\mathcal{U}_\mathcal{Z}} \cdot P_{\mathcal{U}_\mathcal{Y}}\right\| \leq \varepsilon \qquad (3)$$

for all distributions P_X with $H_\infty(X) \geq k$. Here Y is independent of X and uniformly distributed on $\mathcal{Y}$.

The definition implies that $\mathsf{E}(X,y)$ is close to being uniformly distributed on $\mathcal{Z}$ on average over the random choice of $y \leftarrow P_Y$ (cf. (44)). In other words, if X is chosen according to P_X and Y is uniformly distributed and independent of X, then $\mathsf{E}(X,Y)$ is indistinguishable from uniform, even given Y.

In a cryptographic setting, the security of the extracted key $Z := \mathsf{E}(X,Y)$ with respect to an adversary who is given Y is exactly characterized by (3). Indeed, expression (3) quantifies how distinguishable the real system (consisting of (Z,Y)) is from the ideal system, in which Z is uniformly distributed and independent of Y. This is easily generalized to a setting where the adversary is given additional information about X. The additional information can be in the form of a classical random variable (i.e., bits) that is jointly distributed with X or a quantum state (i.e., qubits).

In case the adversary has classical information about X expressed by a random variable E, one can show that this simply reduces the min-entropy of X. If E gives little information about X it follows that even given E and Y, the extracted bits look random. This intuition is made explicit in the following proposition (all proofs in this section can be found in Appendix B).

Proposition 1: Let $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a strong (k,ε)-extractor. Let P_{XE} be a distribution with

$$H_{\mathrm{g}}(X \leftarrow E) \geq k + \log 1/\varepsilon. \qquad (4)$$

Here the guessing-entropy $H_{\mathrm{g}}(X \leftarrow E)$ of X given E is defined as

$$H_{\mathrm{g}}(X \leftarrow E) := -\log \max_{\hat{X}} \Pr[X = \hat{X}] \qquad (5)$$

where the maximum is taken over all random variables $\hat{X}$ such that $X \leftrightarrow E \leftrightarrow \hat{X}$ forms a Markov chain. Then

$$d(\mathsf{E}(X,Y) \leftarrow YE) \leq 2\varepsilon$$

where $P_{YXE} := P_{\mathcal{U}_\mathcal{Y}} \cdot P_{XE}$.

Note that if E is trivial or independent of X the guessing entropy $H_{\mathrm{g}}(X \leftarrow E)$ of X given E is equal to the min-entropy $H_\infty(X)$ of X. The alternative expression

$$H_{\mathrm{g}}(X \leftarrow E) = -\log \mathbb{E}_{e \leftarrow P_E}[\max_x P_{X\,|\,E=e}(x)] \qquad (6)$$

for the guessing entropy shows that it corresponds to a "reasonable" definition of average min-entropy.

Proposition 1 can be applied in the bounded-storage model because the limitation on the adversary's storage implies that his information about X is bounded. More precisely, the guessing probability has the following intuitive property. Any (additional) piece of information W does not increase the success probability in guessing by a significant amount if the size of W is small. More trivially, independent information V does not affect the guessing probability. We express this formally in Lemma 1; versions of this statement are implicit in [1], and more explicitly given in [21].

Lemma 1: Consider a distribution P_{XVWE} with $P_{XV} = P_X \cdot P_V$ and $VW \leftrightarrow X \leftrightarrow E$. Then

$$H_{\mathrm{g}}(X \leftarrow VWE) \geq H_{\mathrm{g}}(X \leftarrow E) - H_0(W).$$

In particular, for every $\varepsilon \geq 0$

$$H_{\mathrm{g}}(X \leftarrow E\,|\,V = v, W = w) \geq H_{\mathrm{g}}(X \leftarrow E) - H_0(W) - \log 1/\varepsilon$$

with probability at least $1 - \varepsilon$ over $(v,w) \leftarrow P_{VW}$.

B. Quantum Adversaries

Let us now discuss the challenge posed by quantum adversaries. Our aim is to show that, similarly as in the classical case, the extracted bits $\mathsf{E}(X,Y)$ are secure even if the adversary is given Y. Such an adversary prepares a quantum state ρ_x on $\mathcal{Q}$ that depends on $X = x$. To obtain maximal information about $\mathsf{E}(X,Y)$, he performs a measurement on his quantum system Q which depends on Y. As a result, his (classical) information E is no longer independent of Y. This means that we cannot view this as merely a reduction of the entropy of the source X. Thus, we cannot directly prove a statement like Lemma 1 when E is replaced by a quantum system Q. In particular, due to the effect of locking [22], we know that there exist short classical keys (Y) that can unlock a lot of classical information (about X) stored in a quantum system Q. In the first part of this paper we will show that if the extractor E extracts a single bit, we can preclude such locking effects (Theorem 1).

Before embarking on this analysis, we point out the following straightforward result. *If* the adversary's measurement does not depend on Y we can essentially apply the classical security proofs. That is, the adversary's measurement produces some classical information E which can be viewed as reducing the

entropy of the source X. If $H_{\min}(X \leftarrow E)$ is still large, then the random variable E does not give much information about X and therefore the extracted bits look random even to such an adversary. This statement is expressed by Proposition 1′ below.

An example of a situation where any measurement outcome E gives a little information about X is the case where the size of the quantum system is sufficiently small. We will express this in a more quantitative form by Proposition 2′ in Section V-A.

Note that we can generalize the guessing-entropy of X given Q to the case where Q is a quantum system. Let $\rho_{XQ} := \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho_x$ be a cq-state. Then

$$H_{\min}(X \leftarrow Q) := -\log \max_{\mathcal{E}} \sum_{x\in\mathcal{X}} P_X(x)\mathrm{tr}(E_x\rho_x) \qquad (7)$$

where the maximum is taken over all POVMs $\mathcal{E} := \{E_x\}_{x\in\mathcal{X}}$ on $\mathcal{Q}$. For a probability distribution P_{XE} with corresponding cc-state ρ_{XE}, definition (7) coincides with the classical definition given in Proposition 1. This is because any POVM with outcome $\hat{X}$ is equivalent to a von Neumann measurement in the computational basis followed by classical post-processing. Thus, the measurement can be seen as a channel $P_{\hat{X}\,|\,E}$ defining a random variable $\hat{X}$ as required.

We now state the nonadaptive quantum version of Proposition 1. It is a direct consequence of the reasoning above.

Proposition 1′: Let $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a strong (k,ε)-extractor, and let $\mathcal{F}$ be a POVM on $\mathcal{Q}$. Then for all cq-states ρ_{XQ} with

$$H_{\min}(X \leftarrow Q) \geq k + \log 1/\varepsilon$$

we have

$$d(\mathsf{E}(X,Y) \leftarrow Y\mathcal{F}(Q)) \leq 2\varepsilon.$$

The reason we are considering a restricted adversary whose measurement does not depend on Y as in Proposition 1′ is not because this is in itself an interesting adversary. In general, an adversary's measurement strategy will depend on Y nontrivially. However, we will see in Section III that the case of a general adversary can be reduced to the type of adversary studied in Proposition 1′ whenever the extractor outputs a single bit.

The guessing entropy used in the statement of Proposition 1′ has properties analogous to the corresponding classical quantity (5). The following is the quantum analogue of Lemma 1 (its proof can again be found in Appendix B). It states that a short additional piece of classical information W does not help much in guessing X if the quantum system Q depends only on X. Again, additional independent information V does not help either.

Lemma 1′: Consider a cccq-state ρ_{XVWQ} with $\rho_{XV} = \rho_X \otimes \rho_V$ and $VW \leftrightarrow X \leftrightarrow Q$. Then

$$H_{\min}(X \leftarrow VWQ) \geq H_{\min}(X \leftarrow Q) - H_0(W)$$

and with probability at least $1-\varepsilon$ over $(v,w) \leftarrow P_{VW}$, we have

$$H_{\min}(X \leftarrow Q\,|\,V=v, W=w) \geq H_{\min}(X \leftarrow Q) - H_0(W) - \log 1/\varepsilon.$$

We now state more precisely what we are aiming to prove about strong extractors. Note that Proposition 1′ only gives a weak security guarantee for the extracted bits $\mathsf{E}(X,Y)$—they are only shown to be secure against an adversary who measures his quantum state *before* receiving Y. To discuss the stronger type of security we aim for, we first state the definition of the nonuniformity in the quantum case.

Definition 3: Let ρ_{ZQ} be an arbitrary cq-state on $\mathcal{Z}\otimes\mathcal{Q}$. The *nonuniformity* $d(Z \leftarrow Q)$ *of* Z *given* Q is defined as

$$d(Z \leftarrow Q) := \|\rho_{ZQ} - \rho_{\mathcal{U}_\mathcal{Z}} \otimes \rho_Q\|$$

where $\rho_{\mathcal{U}_\mathcal{Z}}$ denotes the completely mixed state on $\mathcal{Z}$.

We describe a few basic properties of this definition in Appendix A. In a cryptographic setting, the condition $d(Z \leftarrow Q) \leq \varepsilon$ for some small ε means that the key Z is secure in a setting where Q is controlled by the adversary; as explained in [15] (see also [16], [23]); such a key is, with probability at least $1-\varepsilon$, equivalent to a perfectly secure key.

Defining the security of cryptographic protocols is a very subtle task, especially in the presence of quantum adversaries. The concept of *universal composability* has received much attention in quantum cryptography recently (cf. [15], [23]–[27]). While a general discussion of security definitions and their properties is beyond the scope of this paper, we point out that Definition 3 simply says when a piece of classical information may be regarded as a (universally composable) secure key, given a distributed cq-state. (In the typical scenario involving three parties, it is understood that Alice and Bob both hold Z while Q is controlled by the adversary.) In contrast, security definitions for interactive protocols are generally more involved. We refer the reader to the relevant literature for more details, as our focus is on generating keys which are secure according to Definition 3. A discussion of the merits of this definition can be found in [23].

In the sequel, we aim to show that $d(\mathsf{E}(X,Y) \leftarrow YQ)$ is small for certain strong extractors E and appropriate parameters. This means that the extracted bits are secure even if the adversary is given Y in addition to his quantum system.

In the next section, we will show that for extractors with binary output, the quantity of interest can in fact be bounded by considering an adversary whose strategy does not depend on Y, i.e., he performs a measurement independent of Y as in Proposition 1′. We then use this result in Section IV to construct strong extractors that output several bits.

III. Extractors With Binary Output

We will first sketch the arguments in this section. For an extractor $\mathsf{e} : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ with binary output the nonuniformity of the extracted bit $Z = \mathsf{e}(X,Y)$ given Y and the quantum system Q can be directly related to the success probability in distinguishing two quantum states ρ_0^y and ρ_1^y, for each $y \in \mathcal{Y}$. For a given y, these are the (generally mixed) states of the adversary, conditioned on the extracted bit being 0 or 1, respectively. We modify an argument by Barnum and Knill [28] to bound the optimal success probability in distinguishing ρ_0^y from ρ_1^y for a given y in terms of the success probability resulting from the use of a pretty good measurement [29] $\mathcal{E}_{\mathrm{pgm}}^y$. On the other hand,

we will show that there exists a POVM $\mathcal{F}$ which refines all the pretty good measurements $\{\mathcal{E}^y_{\mathrm{pgm}}\}_{y\in\mathcal{Y}}$ simultaneously; i.e., the outcome of the measurement $\mathcal{E}^y_{\mathrm{pgm}}$ can be obtained by applying $\mathcal{F}$ and classical post-processing. This POVM $\mathcal{F}$ is a pretty good measurement defined by the states $\{\rho_x\}_{x\in\mathcal{X}}$, the conditional states of Q given $X = x$, or, in the bounded-storage model, the states that the adversary prepares upon seeing X. Since the refined measurement $\mathcal{F}$ does not depend on y, we know that it cannot be superior to any classical strategy, see Proposition 1′, and we obtain the main result of this section, Theorem 1.

In the next lemma, we bound the nonuniformity $d(Z \leftarrow Q)$ of a cq-state $\rho_{ZQ} := \sum_{z\in\{0,1\}} p_z|z\rangle\langle z|\otimes\rho_z$ with binary classical part using a pretty good measurement.

Lemma 2: Let $\rho_{ZQ} := \sum_{z\in\{0,1\}} p_z|z\rangle\langle z|\otimes\rho_z$ be a cq-state with binary classical part. Then

$$d(Z\leftarrow Q) \le \sqrt{2d(Z\leftarrow \mathcal{E}_{\mathrm{pgm}}(Q))} + d(Z) \tag{8}$$

where $\mathcal{E}_{\mathrm{pgm}}$ is the pretty good measurement defined by ρ_{ZQ}, i.e., the POVM elements of this measurement are

$$E_z := p_z\rho_Q^{-1/2}\rho_z\rho_Q^{-1/2}, \quad \text{for } z\in\{0,1\}.$$

Proof: By definition

$$d(Z\leftarrow Q) = \sum_{z=0}^{1}\left\| p_z\rho_z - \frac{1}{2}\rho_Q\right\| = \|p_0\rho_0 - p_1\rho_1\|. \tag{9}$$

Let $\Delta := p_0\rho_0 - p_1\rho_1$ and let $\Delta =: A^+ - A^-$ with $A^+ \ge 0$, $A^- \ge 0$ be the decomposition of Δ into a nonnegative and a negative part. Then

$$\begin{aligned}\|p_0\rho_0 - p_1\rho_1\| &= \frac{1}{2}(\mathsf{tr}(A^+) + \mathsf{tr}(A^-))\\ &= \mathsf{tr}(A^+) - \frac{1}{2}\mathsf{tr}(\Delta)\\ &= \mathsf{tr}(\mathbb{P}\Delta) - \frac{1}{2}(p_0 - p_1)\end{aligned}$$

where $\mathbb{P}$ is the projector onto the support of A^+. We will do some work to show that

$$\mathsf{tr}(\mathbb{P}\Delta) \le \sqrt{2d(Z\leftarrow\mathcal{E}_{\mathrm{pgm}}(Q))} \tag{10}$$

where $\mathcal{E}_{\mathrm{pgm}}$ is the pretty good measurement that distinguishes ρ_1 and ρ_1. By noting that

$$-\frac{1}{2}(p_0 - p_1) \le \frac{1}{2}|p_0 - p_1| = d(Z),$$

we obtain the desired result, (8). Consider thus the quantity $\mathsf{tr}(\mathbb{P}\Delta)$. We can bound

$$\mathsf{tr}(\mathbb{P}\Delta) \le \sqrt{\mathsf{tr}(A^\dagger A)\mathsf{tr}(B^\dagger B)} \tag{11}$$

by applying the operator Cauchy–Schwarz inequality to the operators

$$\begin{aligned}A &:= \rho_Q^{1/4}\mathbb{P}\rho_Q^{1/4}\\ B &:= \rho_Q^{-1/4}\Delta\rho_Q^{-1/4}.\end{aligned}$$

However

$$\begin{aligned}\mathsf{tr}(A^\dagger A) &= \mathsf{tr}(\rho_Q^{1/2}\mathbb{P}\rho_Q^{1/2}\mathbb{P})\\ &\le \mathsf{tr}(\rho_Q^{1/2}\mathbb{P}\rho_Q^{1/2})\\ &\le \mathsf{tr}(\rho_Q) = 1\end{aligned} \tag{12}$$

where we used the fact that $\mathbb{P}\le 1$ and the fact that $\rho_Q^{1/2}\mathbb{P}\rho_Q^{1/2}$ is nonnegative. On the other hand, by the definition of the pretty good measurement $\mathcal{E}_{\mathrm{pgm}} = \{E_0, E_1\}$ we have

$$\begin{aligned}\mathsf{tr}(B^\dagger B) &= \mathsf{tr}(\rho_Q^{-1/2}\Delta\rho_Q^{-1/2}\Delta)\\ &= \mathsf{tr}(E_0\Delta) - \mathsf{tr}(E_1\Delta)\\ &= P_{\mathrm{succ}}(\mathcal{E}_{\mathrm{pgm}}) - p_1\mathsf{tr}(E_0\rho_1) - p_0\mathsf{tr}(E_1\rho_0)\\ &= 2P_{\mathrm{succ}}(\mathcal{E}_{\mathrm{pgm}}) - 1.\end{aligned} \tag{13}$$

Here we have used the definition of the success probability $P_{\mathrm{succ}}(\{E_0,E_1\}) := p_0\mathsf{tr}(E_0\rho_0) + p_1\mathsf{tr}(E_1\rho_1)$, and the fact that $E_0 + E_1 = 1$ and $p_0 + p_1 = 1$ in the last step. Note that the probability of success $P_{\mathrm{succ}}(\mathcal{E})$ for a fixed POVM $\mathcal{E}$ is the same as the probability of successfully distinguishing an instance drawn from the distribution of measurement outcomes when applying $\mathcal{E}$ to ρ_0 and ρ_1, respectively, with *a priori* probabilities p_0 and p_1. The latter task corresponds to the binary decision problem with states (or random variables) $\mathcal{E}(\rho_0)$, $\mathcal{E}(\rho_1)$, and priors p_0, p_1. Now we invoke Helstrom's theorem [30] which says that the success probability of distinguishing two quantum states σ_0 and σ_1 with priors p_0 and p_1 using an optimal POVM $\mathcal{E}_{\mathrm{opt}}$ is equal to $P_{\mathrm{succ}}(\mathcal{E}_{\mathrm{opt}}) = \frac{1}{2} + \|p_0\sigma_0 - p_1\sigma_1\|$. We apply this theorem for $\sigma_z = \mathcal{E}_{\mathrm{pgm}}(\rho_z)$ and write

$$P_{\mathrm{succ}}(\mathcal{E}_{\mathrm{pgm}}) = \frac{1}{2} + d(Z\leftarrow\mathcal{E}_{\mathrm{pgm}}(Q))$$

(cf. (9)). Combining this with (11), (12), and (13) yields (10), as desired. □

Now the goal is to bound the nonuniformity $d(\mathsf{e}(X,Y)\leftarrow YQ)$ for extractors e with binary output when Q is a quantum system which depends on X. For this we consider the cccq-state $\rho_{ZXYQ} \equiv \rho_{\mathsf{e}(X,Y)XYQ}$ which has the form

$$\rho_{ZXYQ} = \sum_{x,y,z=\mathsf{e}(x,y)} P_X(x)P_Y(y)|zxy\rangle\langle zxy|\otimes\rho_x \tag{14}$$

where $P_Y(y) = \frac{1}{|\mathcal{Y}|}$ for every $y\in\mathcal{Y}$. For this state, one can express the nonuniformity $d(Z\leftarrow YQ)$ as (cf. (2))

$$\mathbb{E}_{y\leftarrow P_Y}\left[\sum_{z\in\{0,1\}}\left\|\sum_x P_{Z\,|\,X=x,Y=y}(z)P_X(x)\rho_x - \frac{1}{2}\rho_Q\right\|\right]. \tag{15}$$

Note that $P_{Z\,|\,X=x,Y=y}(z)$ is 1 or 0, depending on whether or not $\mathsf{e}(x,y) = z$. It is straightforward to verify that

$$\begin{aligned}p_z^y\rho_z^y &:= \sum_{x\in\mathcal{X}} P_{Z\,|\,Y=y}(z)P_{X\,|\,Y=y,Z=z}(x)\rho_x\\ &= \sum_{x\in\mathcal{X}} P_{Z\,|\,X=x,Y=y}(z)P_X(x)\rho_x\end{aligned} \tag{16}$$

where we introduced for each $y \in \mathcal{Y}$ and $z \in \{0,1\}$ the density matrix

$$\rho_z^y := \sum_{x \in \mathcal{X}} P_{X \mid Y=y, Z=z}(x) \rho_x \tag{17}$$

with the normalising factor

$$p_z^y := P_{Z \mid Y=y}(z) = \Pr_{x \leftarrow X}[\mathsf{e}(x,y) = z]. \tag{18}$$

The state ρ_z^y is the state of Q conditioned on $\mathsf{e}(X,y) = z$; for any given $y \in \mathcal{Y}$, the two states ρ_0^y and ρ_1^y have *a priori* probabilities p_z^y with $z \in \{0,1\}$. From this definition, it is clear that

$$\sum_z p_z^y \rho_z^y = \sum_x P_X(x) \rho_x = \rho_Q \tag{19}$$

which is *independent* of y. This observation will be essential in the proof of the following theorem.

Applying Helstrom's theorem gives an intuitive interpretation of the quantity of interest (which we state, but do not need later in the proof): $\frac{1}{2} + d(\mathsf{e}(X,Y) \leftarrow YQ)$ is the maximal average success probability when distinguishing ρ_0^y and ρ_1^y with *a priori* probabilities p_0^y and p_1^y, over random $y \leftarrow P_Y$. This follows by combining (19) with (15) and (17), (18).

We are ready to derive the main result of this section:

Theorem 1: Let $\mathsf{e} : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a strong (k,ε)-extractor. Then for all ρ_{XQ} with

$$H_{\mathsf{g}}(X \leftarrow Q) \geq k + \log 1/\varepsilon$$

we have

$$d(\mathsf{e}(X,Y) \leftarrow YQ) \leq 3\sqrt{\varepsilon}$$

where $\rho_{YXQ} := \rho_{\mathcal{U}_\mathcal{Y}} \otimes \rho_{XQ}$.

Proof: By (2) (cf. (44)) we can express

$$d(\mathsf{e}(X,Y) \leftarrow YQ) = \mathbb{E}_{y \leftarrow P_Y}[d(\mathsf{e}(X,y) \leftarrow Q)]. \tag{20}$$

We can apply the pretty good measurement bound of Lemma 2 for each $y \in \mathcal{Y}$ to the state $\rho_{\mathsf{e}(X,y)Q} = \sum_{z \in \{0,1\}} p_z^y |z\rangle\langle z| \otimes \rho_z^y$, where the density matrices ρ_z^y and their associated probabilities p_z^y are defined in (17) and (18). We get

$$d(\mathsf{e}(X,y) \leftarrow Q) \leq \sqrt{2 d(\mathsf{e}(X,y) \leftarrow \mathcal{E}_{\mathrm{pgm}}^y(Q))} + d(\mathsf{e}(X,y))$$

for every $y \in \mathcal{Y}$. Taking the expectation over $y \leftarrow P_Y$ again and using the convexity of the square root gives

$$d(\mathsf{e}(X,Y) \leftarrow YQ) \leq \sqrt{\mathbb{E}_{y \leftarrow P_Y}\left[2 d(\mathsf{e}(X,y) \leftarrow \mathcal{E}_{\mathrm{pgm}}^y(Q))\right]} + d(\mathsf{e}(X,Y) \leftarrow Y) \tag{21}$$

by (2) (see also (43)). Since

$$H_\infty(X) \geq H_{\mathsf{g}}(X \leftarrow Q)$$

the second term in (21) is upper-bounded by ε. Let us now consider the details of the pretty good measurement $\mathcal{E}_{\mathrm{pgm}}^y$. The measurement $\mathcal{E}_{\mathrm{pgm}}^y = \{E_z^y\}_{z \in \{0,1\}}$ is determined by the POVM elements

$$E_z^y := p_z^y (G^y)^{-1/2} \rho_z^y (G^y)^{-1/2} \tag{22}$$

where, as argued above (19)

$$G^y = \sum_{z \in \mathcal{Z}} p_z^y \rho_z^y = \rho_Q \tag{23}$$

is independent of y. This fact allows us to define a new pretty good measurement $\mathcal{F}$ which does not depend on y, but is equally good or better in estimating Z from Q and Y. This new pretty good measurement $\mathcal{F} = \{F_x\}_{x \in \mathcal{X}}$ has POVM elements

$$F_x := P_X(x) \rho_Q^{-1/2} \rho_x \rho_Q^{-1/2}.$$

Expressed differently, $\mathcal{F}$ is simply the pretty good measurement defined by the ensemble $\{P_X(x), \rho_x\}$. From (16), (22), and (23) above one can see that

$$E_z^y = \sum_{x \in \mathcal{X}} P_{Z \mid X=x, Y=y}(z) F_x. \tag{24}$$

In other words, the results of the measurements $\{\mathcal{E}_{\mathrm{pgm}}^y\}_{y \in \mathcal{Y}}$ can in fact be obtained by first estimating x by measuring the quantum system Q with $\mathcal{F} = \{F_x\}_{x \in \mathcal{X}}$. Then we infer z for a given y by computing $z = \mathsf{e}(x,y)$. On a more technical level, one needs to show that for every $y \in \mathcal{Y}$, the nonuniformity given the measurement outcome of the measurement $\mathcal{E}_{\mathrm{pgm}}^y$ is smaller than or equal to the nonuniformity given the outcome of the refined measurement $\mathcal{F}$. We have summarized these technical details in Lemma 6 proved in Appendix C. Formally, we have

$$d(\mathsf{e}(X,y) \leftarrow \mathcal{E}_{\mathrm{pgm}}^y(Q)) \leq d(\mathsf{e}(X,y) \leftarrow \mathcal{F}(Q)).$$

Taking the expectation over $y \leftarrow P_Y$ gives (cf. (44))

$$\mathbb{E}_{y \leftarrow P_Y}\left[d\left(\mathsf{e}(X,y) \leftarrow \mathcal{E}_{\mathrm{pgm}}^y(Q)\right)\right] \leq d(\mathsf{e}(X,Y) \leftarrow Y\mathcal{F}(Q)).$$

Since $\mathcal{F}$ does not depend on y we have reduced our problem to the simple scenario where the quantum system is measured before the adversary obtains y. Thus, we can apply Proposition $1'$

$$d(\mathsf{e}(X,Y) \leftarrow Y\mathcal{F}(Q)) \leq 2\varepsilon. \tag{25}$$

We conclude with (21) that

$$d(\mathsf{e}(X,Y) \leftarrow YQ) \leq 2\sqrt{\varepsilon} + \varepsilon$$

hence, the claim follows. □

We point out that the proof of Theorem 1 reveals that

$$\max_{\{\rho_x\}_x} d(\mathsf{e}(X,Y) \leftarrow YQ) \leq 3\sqrt{\max_{P_{E \mid X}} d(\mathsf{e}(X,Y) \leftarrow YE)}.$$

In this inequality, the maximums are over all families of conditional states $\{\rho_x = \rho_{Q \mid X=x}\}_{x \in \mathcal{X}}$ and conditional distributions $P_{E \mid X}$ defining states ρ_{XQ} and distributions P_{XE}, respectively, with

$$H_{\mathsf{g}}(X \leftarrow Q), H_{\mathsf{g}}(X \leftarrow E) \geq k + \log 1/\varepsilon.$$

This shows that the security of a single extracted bit is only quadratically worse with respect to an adversary that has a similar amount of quantum instead of classical information.

We now show that even if the adversary is given additional information V which is independent of X and a short bit string W which might depend on X, the extracted bit looks secure. This statement will be used below to prove that certain extractors which output several bits can also safely be used in a cryptographic context (cf. Theorem 2).

Corollary 1: Let $\mathsf{e} : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a strong (k,ε)-extractor. Let ρ_{XVWQ} be a cccq-state with $\rho_{XV} = \rho_X \otimes \rho_V$, $VW \leftrightarrow X \leftrightarrow Q$, and

$$H_{\mathsf{g}}(X \leftarrow Q) \geq k + H_0(W) + 2\log 1/\varepsilon.$$

Then

$$d(\mathsf{e}(X,Y) \leftarrow YVWQ) \leq 4\sqrt{\varepsilon}$$

where $\rho_{YXVWQ} := \rho_{\mathcal{U}_\mathcal{Y}} \otimes \rho_{XVWQ}$.

Proof: Let $\alpha := d(\mathsf{e}(X,Y) \leftarrow YVWQ)$ be the quantity of interest. Then by (2) (see also (43))

$$\alpha = \mathbb{E}_{(v,w)\leftarrow P_{VW}}[d(\mathsf{e}(X,Y) \leftarrow YQ \,|\, V=v, W=w)]$$

where the term in brackets is the nonuniformity of $\mathsf{e}(X,Y)$ with respect to the conditional state $\rho_{XQ\,|\,V=v,W=w}$. By Lemma $1'$, we have

$$H_{\mathsf{g}}(X \leftarrow Q \,|\, V=v, W=w) \geq k + \log 1/\varepsilon \qquad (26)$$

with probability at least $1-\varepsilon$ over random $(v,w) \leftarrow P_{VW}$. For any (v,w) for which (26) is satisfied, we have

$$d(\mathsf{e}(X,Y) \leftarrow YQ \,|\, V=v, W=w) \leq 3\sqrt{\varepsilon}$$

by Theorem 1. Thus

$$\alpha \leq 3\sqrt{\varepsilon} + \varepsilon,$$

and the claim follows. □

IV. Extractors With Nonbinary Output

In this section, we will consider strong extractors which output several bits. We first show how to use independent seeds $y_1,\ldots,y_m$ to extract m bits. The security of the extracted bits in the quantum setting will follow from applying our bound for binary extractors, Theorem 1, in combination with a quantum version of the so-called hybrid argument. By a similar technique, we will show how to extract more bits under stronger assumptions. Let us first discuss the hybrid argument.

Consider a cq-state of the form ρ_{ZQ}, where $Z = (Z_1,\ldots,Z_m)$ is an m-bit string. We aim to find a bound on $d(Z \leftarrow Q)$ in terms of nonuniformities of binary random variables.

By definition, we have

$$d(Z \leftarrow Q) = \left\| \rho_{ZQ} - \rho_{\mathcal{U}_{\{0,1\}}}^{\otimes m} \otimes \rho_Q \right\|.$$

Let us define for $i = 0,\ldots,m$ the states

$$\rho^{(i)} := \rho_{\mathcal{U}_{\{0,1\}}}^{\otimes m-i} \otimes \rho_{Z^iQ}$$

on $\{0,1\}^m \otimes \mathcal{Q}$, where we use the abbreviation $z^i := (z_1,\ldots,z_i)$ to refer to the first i bits of $z \in \{0,1\}^m$. Clearly, we have $\rho^{(m)} = \rho_{ZQ}$ and $\rho^{(0)} = \rho_{\mathcal{U}_{\{0,1\}}}^{\otimes m} \otimes \rho_Q$. We use the "telescoping" sum

$$\rho^{(0)} - \rho^{(m)} = \sum_{i=0}^{m-1} \rho^{(i)} - \rho^{(i+1)}$$

which, by the triangle inequality, implies that

$$d(Z \leftarrow Q) \leq \sum_{i=0}^{m-1} \left\| \rho^{(i+1)} - \rho^{(i)} \right\|.$$

But

$$\begin{aligned} \left\| \rho^{(i+1)} - \rho^{(i)} \right\| &= \left\| \rho_{\mathcal{U}_{\{0,1\}}}^{\otimes m-i-1} \otimes \rho_{Z^{i+1}Q} - \rho_{\mathcal{U}_{\{0,1\}}}^{\otimes m-i} \otimes \rho_{Z^iQ} \right\| \\ &= \left\| \rho_{Z^{i+1}Q} - \rho_{\mathcal{U}_{\{0,1\}}} \otimes \rho_{Z^iQ} \right\| \\ &= \left\| \rho_{Z_{i+1}Z^iQ} - \rho_{\mathcal{U}_{\{0,1\}}} \otimes \rho_{Z^iQ} \right\|. \end{aligned}$$

We thus arrive at the following conclusion:

$$d(Z \leftarrow Q) \leq \sum_{i=0}^{m-1} d(Z_{i+1} \leftarrow Z^iQ) \qquad (27)$$

where $Z^i = (Z_1,\ldots,Z_i)$.

Let us now state and prove the main theorem.

Theorem 2: Let $\mathsf{e} : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a strong (k,ε)-extractor, and let

$$\begin{aligned} \mathsf{E}^m : \mathcal{X} \times \mathcal{Y}^m &\to \{0,1\}^m \\ (x, y_1,\ldots,y_m) &\mapsto (\mathsf{e}(x,y_1),\ldots,\mathsf{e}(x,y_m)). \end{aligned}$$

Then for all cq-states ρ_{XQ} with

$$H_{\mathsf{g}}(X \leftarrow Q) \geq k + m + 2\log 1/\varepsilon \qquad (28)$$

we have

$$d(\mathsf{E}^m(X,Y^m) \leftarrow Y^mQ) \leq 4m\sqrt{\varepsilon}$$

where $\rho_{Y^mXQ} := \rho_{\mathcal{U}_{\mathcal{Y}^m}} \otimes \rho_{XQ}$.

Proof: We use (27) to get

$$d(\mathsf{E}^m(X,Y^m) \leftarrow Y^mQ) \leq \sum_{i=0}^{m-1} d(Z_{i+1} \leftarrow Z^iY^mQ) \qquad (29)$$

where $Z^m = \mathsf{E}^m(X,Y^m)$. Observe that $(Y_{i+2},\ldots,Y_m)$ is independent of $Z^{i+1}Y^{i+1}Q$, which by (45) gives

$$d(Z_{i+1} \leftarrow Z^iY^mQ) = d(Z_{i+1} \leftarrow Z^iY^{i+1}Q). \qquad (30)$$

But

$$\begin{aligned} d(Z_{i+1} \leftarrow Z^iY^{i+1}Q) &= d(Z_{i+1} \leftarrow Y_{i+1}Z^iY^iQ) \\ &= d(\mathsf{e}(X,\tilde{Y}) \leftarrow \tilde{Y}\,\mathsf{E}^i(X,Y^i)Y^iQ) \end{aligned}$$

where $\tilde{Y} \equiv Y_{i+1}$. Applying Corollary 1 to $(V,W) = (Y^i, \mathsf{E}^i(X,Y^i))$ yields

$$d(Z_{i+1} \leftarrow Z^i Y^{i+1} Q) \leq 4\sqrt{\varepsilon} \tag{31}$$

for every $i = 0,\ldots,m-1$. We have made use of the fact that $H_0(W) = H_0(\mathsf{E}^i(X,Y^i)) \leq m$ by definition. The claim then follows from (29), (30), and (31). □

In the next section, we study the implications of Theorem 2 for the bounded-storage model. We will see that the bound on the storage of the adversary translates into an upper bound on the guessing probability, as required (cf. (28)). We will then give a concrete example of an extractor for the bounded-storage model with quantum adversaries.

Before continuing, however, let us point out that in certain situations, we can use the hybrid argument to show that the seed Y can be reused several times. This gives more efficient randomness extractors (under stronger assumptions about the inital cq-state ρ_{XQ}). Following similar terminology in the literature on extractors, we introduce the following notion.

Definition 4: A cq-state ρ_{XQ} where $X = (X_1,\ldots,X_m)$ consists of m parts is a *k-blockwise state* if for all $i = 0,\ldots,m-1$

$$H_{\mathsf{g}}(X_{i+1} \leftarrow X^i Q) \geq k.$$

We will now show how to extract multiple bits from such a cq-state by reusing the seed. This is interesting for several reasons. First, k-blockwise states arise naturally in realistic situations such as the bounded-storage model. We will discuss this in more detail below (cf. Section V-C). Second, extractors for k-blockwise probability distributions are often used to construct (classical) extractors by transforming the input distribution to a k-blockwise distribution. It might therefore be possible to obtain extractor constructions for the quantum case using similar lines of reasoning.

Theorem 3: Let $\mathsf{e} : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a strong (k,ε)-extractor, and let

$$\tilde{\mathsf{E}}^L : \mathcal{X}^L \times \mathcal{Y}^m \to \{0,1\}^{Lm}$$
$$(x_1,\ldots,x_L,y) \mapsto (\mathsf{E}^m(x_1,y),\ldots,\mathsf{E}^m(x_L,y))$$

where $\mathsf{E}^m : \mathcal{X} \times \mathcal{Y}^m \to \{0,1\}^m$ is defined as in Theorem 2. Then

$$d(\tilde{\mathsf{E}}^L(X^L,Y) \leftarrow YQ) \leq 4Lm\sqrt{\varepsilon}$$

for all $(k+m+2\log 1/\varepsilon)$-blockwise states ρ_{XQ} on $\mathcal{X}^L \otimes \mathcal{Q}$, where $\rho_{YXQ} := \rho_{\mathcal{U}_\mathcal{Y}} \otimes \rho_{XQ}$.

Proof: With (27) we get

$$d(\tilde{\mathsf{E}}^L(X^L,Y) \leftarrow YQ) \leq \sum_{i=0}^{L-1} d(Z_{i+1} \leftarrow Z^i YQ) \tag{32}$$

where $Z^L \equiv \tilde{\mathsf{E}}^L(X^L,Y)$. Since (Z^i,Y) is a function of (X^i,Y) and since applying functions does not increase the trace distance, we obtain

$$d(\tilde{\mathsf{E}}^L(X^L,Y) \leftarrow YQ) \leq \sum_{i=0}^{L-1} d(Z_{i+1} \leftarrow X^i YQ). \tag{33}$$

But $d(Z_{i+1} \leftarrow X^i YQ) = d(\mathsf{E}^m(X_{i+1},Y) \leftarrow YX^iQ)$, and $\rho_{YX_{i+1}X^iQ} = \rho_{\mathcal{U}_\mathcal{Y}} \otimes \rho_{X_{i+1}X^iQ}$. Moreover

$$H_{\mathsf{g}}(X_{i+1} \leftarrow X^i Q) \geq k + m + 2\log 1/\varepsilon.$$

by assumption. Thus, we can apply Theorem 2 and the claim follows. □

V. The Bounded-Storage Model With a Quantum Adversary

A. Bounded Storage, Guessing Entropy, and Extractors

In the classical version of the bounded-storage model, the security of the extracted bits is a direct consequence of the property of the extractor given in Proposition 1 and the fact that an adversary has limited information about X. The latter fact is expressed by the following well-known proposition, whose proof we omit, as it is trivial. It states that an adversary who has $H_0(E)$ bits of storage cannot predict X well, and is also known as "chain-rule for min-entropy."

Proposition 2: Let P_{XE} be an arbitrary distribution. Then

$$H_{\mathsf{g}}(X \leftarrow E) \geq H_\infty(X) - H_0(E).$$

Together with Lemma 1, it follows that a strong (k,ϵ)-extractor has the property that $d(\mathsf{e}(X,Y) \leftarrow YE) \leq 2\epsilon$ for all P_{XE} with

$$H_\infty(X) \geq k + H_0(E) + \log 1/\varepsilon.$$

Thus, the security of the extracted key can be directly derived from the strong extractor property and the bounded-storage assumption. The main challenge is to construct strong extractors which satisfy all the additional requirements for applicability in the bounded-storage model (see Section V-B).

What about quantum storage? We show that a similar reasoning applies; given an extractor which is characterized by the guessing-entropy $H_{\mathsf{g}}(X \leftarrow Q)$, the storage bound can be translated into a security guarantee. We first show that X cannot be guessed by measuring Q when the number of qubits constituting Q is limited.

Proposition 2′: Let ρ_{XQ} be a cq-state. Then

$$H_{\mathsf{g}}(X \leftarrow Q) \geq H_\infty(X) - H_0(Q).$$

Proof: Consider a POVM $\mathcal{E} := \{E_x\}_{x\in\mathcal{X}}$ on $\mathcal{Q}$ that maximizes the expression defining $H_{\mathsf{g}}(X \leftarrow Q)$ (cf. (7)). Then

$$\begin{aligned} 2^{-H_{\mathsf{g}}(X\leftarrow Q)} &= \sum_x P_X(x)\mathsf{tr}(E_x\rho_x) \\ &= \mathsf{tr}\left(\left(\sum_x |x\rangle\langle x| \otimes E_x\right)\rho_{XQ}\right) \\ &\leq 2^{-H_\infty(XQ)}\mathsf{tr}\left(\sum_x |x\rangle\langle x| \otimes E_x\right). \end{aligned}$$

The statement then follows from the fact that

$$\mathsf{tr}\left(\sum_x |x\rangle\langle x| \otimes E_x\right) = \mathsf{tr}\left(\sum_x E_x\right) = \mathsf{tr}(1_\mathcal{Q}) = 2^{H_0(Q)}.$$

□

By combining Proposition 2′ with Theorem 2, we obtain a way of constructing strong extractors for the bounded-storage model in the presence of quantum adversaries: the statement of Theorem 2 holds when (28) is replaced by the stronger condition

$$H_\infty(X) \geq k + m + H_0(Q) + 2\log 1/\varepsilon. \tag{34}$$

Before applying this result to obtain a concrete construction, let us elaborate on a recent example which shows that not every strong extractor yields secure bits in the bounded-quantum-storage model.

Remark 1: Gavinsky, Kempe , and de Wolf [20] consider the function

$$\begin{aligned} \mathsf{e} : \{0,1\}^n \times \omega_n &\rightarrow \{0,1\} \\ ((x_1,\ldots,x_n),\{y_1,y_2\}) &\mapsto x_{y_1} \oplus x_{y_2} \end{aligned}$$

where $\oplus$ denotes bitwise addition modulo 2 and where ω_n is the set of pairs (y_1, y_2) of distinct indices $y_1, y_2 \in \{1,\ldots,n\}$. They then study the function E^m restricted to the set $\{0,1\}^n \times \Omega_m$, where $\Omega_m \subset \omega_n^m$ is the subset of disjoint m-tuples. Let us call this restriction $\tilde{\mathsf{E}}_m$ and let $\tilde{Y}_m$ be uniform on Ω_m. In our terminology, they show the following. There is an $\alpha \approx 1/\sqrt{\log n}$ such that for large enough n and $m := \alpha n$, the quantity $d(\tilde{\mathsf{E}}_m(X,\tilde{Y}_m) \leftarrow \tilde{Y}_m E)$ is small for any classical random variable E with $H_0(E) \leq \sqrt{n}$, whereas $d(\tilde{\mathsf{E}}_m(X,\tilde{Y}_m) \leftarrow \tilde{Y}_m Q)$ is large if Q is quantum and $H_0(Q)$ is polylogarithmic in n.

This statement does not contradict Theorem 2 which cannot be applied in this situation. This is because the function $\tilde{E}_m$ does not have the required form. While Theorem 1 tells us that the difference between classical and quantum prior information is limited in the case of extractors with binary output, this example shows that the case of general extractors which output several bits is more subtle.

B. Extractors for the Bounded-Storage Model: An Explicit Example

In this subsection, we give a concrete example of a function $\mathsf{E} : \{0,1\}^n \times \{0,1\}^t \rightarrow \{0,1\}^m$ which can be used in the bounded-storage model in the presence of a quantum adversary. Let us first discuss what additional requirements such a function has to satisfy.

Typical parameters of the bounded-storage model are as follows: For some $1 \geq \alpha > \beta > 0$, $H_\infty(X) \geq \alpha n$ and $H_0(Q) \leq \beta n$. Here, the parameter α is called the *min-entropy rate*, whereas β is referred to as the *storage rate*. The amount of memory available to the honest parties, Alice and Bob, on the other hand, is supposed to be much more limited. Typically, it is assumed that they have only $O(\log n)$ bits of storage. Expressed differently, the scheme should be secure even if the adversary is significantly more powerful than the participating honest parties.

The fact that Alice and Bob have only $O(\log n)$ bits of memory implies that the strong extractor must have seed length $\log|\mathcal{Y}| = t$ of that order. Moreover, the extractor has to be (efficiently) computable with limited memory. This is the case if E is *ℓ-local*, meaning that it only depends on a small number ℓ (instead of n) physical bits of its first argument, where the ℓ bit locations are determined by the second argument. Note that a different solution to the latter problem was suggested by Lu [11], who considers so-called on-line computable functions.

Due to these requirements, finding explicit, efficiently computable constructions for the bounded-storage model is a rather intricate problem, which has been studied for some time [8]–[12]. Here we consider a construction by Vadhan. By choosing the output to be a single bit, Theorem 8.5 in [12] gives an ℓ-local strong (k,ε)-extractor $\mathsf{e} : \{0,1\}^n \times \{0,1\}^t \rightarrow \{0,1\}$ with

$$t = \log n + O(\log 1/\varepsilon) \tag{35}$$

$$\ell = \frac{3}{2}\frac{n}{k} + O(\log 1/\varepsilon) \tag{36}$$

for every $\varepsilon > \exp(-n/2^{O(\log^* n)})$. (We set $\kappa = 1/2$ in [12, Theorem 8.5] for convenience—this parameter controls the number ℓ of randomizer bits read.) The term $\log^* n$ refers to the iterated logarithm of n, which is defined recursively as $\log^* n = 1 + \log^*(\log n)$ for $n > 1$ and 0 otherwise.

Suppose we want to achieve an error ε, using Theorem 2. Then the error for the one-bit-extractor e must be upper-bounded by $(\varepsilon/4m)^2$. Inserting this into (35) and (36) gives the following.

Corollary 2: For any $\varepsilon > 4m\exp(-n/2^{O(\log^* n)})$, there is an ℓ-local function

$$\mathsf{E} : \{0,1\}^n \times \{0,1\}^t \rightarrow \{0,1\}^m$$

with

$$\begin{aligned} t &= m\log n + O(m\log m + m\log 1/\varepsilon) \\ \ell &= \frac{3}{2}\frac{nm}{k - m - 2\log 1/\varepsilon} + O(m\log m + m\log 1/\varepsilon) \end{aligned}$$

such that $d(\mathsf{E}(X,Y) \leftarrow YQ) \leq \varepsilon$ for all ρ_{XQ} with

$$H_\infty(X) - H_0(Q) \geq k$$

where $\rho_{YXQ} = \rho_{\mathcal{U}_{\{0,1\}^t}} \otimes \rho_{XQ}$.

In terms of the min-entropy rate α and the storage-rate β, our result implies that for any $\alpha > \beta$, there is an extractor which uses $O(m\log n + m\log 1/\varepsilon)$ bits of initial key, outputs m bits with security ε, and reads $O(m\log m + m\log 1/\varepsilon)$ bits from the randomizer X. (This follows under the assumption $m \leq \gamma n$ for some constant $\gamma < \alpha$—we typically have $m = O(\log n)$ in the bounded-storage model.) In comparison, the best known classical construction [12] uses $O(\log n + \log 1/\varepsilon)$ bits of key and reads $O(m + \log 1/\varepsilon)$ from X.

Note that the extractor of Corollary 2 outputs fewer bits than the number of initial key bits t it uses. We stress that it still achieves key expansion in the bounded storage model, even if the output is chosen to be a single bit ($m = 1$). This is because the pair $\mathsf{E}(X,Y)Y$ is a universally composable key. The construction of [10] for the bounded storage model with a classical adversary is the first which achieves significant key expansion (i.e., $m \gg t$). It is an open problem to find strong extractors which output more secure bits than the number of key bits

consumed in the presence of prior quantum information. In the next subsection, we show how to extract more bits under slightly stronger assumptions.

C. Independent Randomizers

In the so-called satellite scenario [8], the randomizer X is assumed to consist of a sequence of random bits that are publicly broadcast in sequence. In this situation, it is clear that if we partition X into blocks $X = (X_1, \ldots, X_L)$, the random variables corresponding to the blocks are independent. What is more interesting is that if the adversary is allowed to prepare a quantum system Q adaptively, the resulting cq-state ρ_{XQ} is a k-blockwise state. This is a consequence of the fact that taking the previous blocks X^i into account when storing and retrieving information about X_{i+1} does not help the adversary if X_{i+1} is independent of X^i. We can express this formally by the following result, with the set $\mathcal{S}$ corresponding to all states on a Hilbert space of limited dimension in the bounded-storage model.

Lemma 3: Let $P_{XX'} = P_X \cdot P_{X'}$ be a probability distribution of independent random variables and let $\mathcal{S}$ be a set of states. Then

$$\min_{\rho_{XX'Q}} H_{\min}(X \leftarrow X'Q) \geq \min_{\rho_{XQ}} H_{\min}(X \leftarrow Q) \tag{37}$$

where the minima are over all states of the form

$$\rho_{XX'Q} = \sum_{x,x'} P_{XX'}(x,x') |x\rangle\langle x| \otimes |x'\rangle\langle x'| \otimes \rho_x^{x'}$$

with $\rho_x^{x'} \in \mathcal{S}$ and

$$\rho_{XQ} = \sum_x P_X(x) |x\rangle\langle x| \otimes \rho_x, \quad \rho_x \in \mathcal{S}$$

respectively.

Proof: Let $\{\rho_x^{x'}\}_{x,x'}$ be a family of states such that the corresponding state $\rho_{XX'Q}$ achieves the minimum on the left-hand side (LHS) of (37). Then

$$2^{-H_{\min}(X \leftarrow X'Q)} = \mathbb{E}_{x' \leftarrow P_{X'}} \left[2^{-H_{\min}(X \leftarrow Q \,|\, X'=x')} \right].$$

However

$$\begin{aligned} 2^{-H_{\min}(X \leftarrow Q \,|\, X'=x')} &= \max_{\{E_x\}_x} \sum_{x \in \mathcal{X}} P_X(x) \mathrm{tr}\left(E_x \rho_x^{x'}\right) \\ &= 2^{-H_{\min}(X \leftarrow Q)} \end{aligned}$$

where the latter expression denotes the guessing entropy of X given Q in the state

$$\rho_{XQ}^{x'} = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x| \otimes \rho_x^{x'}.$$

The claim directly follows from this. □

If the randomizer X consists of several independent parts $X = (X_1, \ldots, X_L)$ which satisfy $H_\infty(X_i) \geq H_0(Q) + k$ for all i, we can therefore use our hybrid construction (Theorem 3) in conjunction with Corollary 2. As an example, consider the case where each of the blocks X_i consists of n bits with min-entropy rate α. We then obtain an extractor $\mathsf{E} : \{0,1\}^{Ln} \times \{0,1\}^t \to \{0,1\}^{Lm}$ which uses $t = mO(\log n + \log L + \log 1/\varepsilon)$ bits of initial key, reads $mO(\log m + \log L + \log 1/\varepsilon)$ bits from X and gives an ε-secure output in the presence of an adversary with storage rate $\beta < \alpha$. In particular, this construction can extend the key of the honest parties by more than the number of initial key bits. This implies that Alice and Bob end up with a longer key even if the adversary later learns the initial key Y.

VI. Tomography-Based Approach to General Extractors

The results of Section III imply that the security of a single extracted bit is similar with respect to an adversary that has quantum instead of classical resources.

This is not true for general extractors which output several bits, as shown in [20] by an explicit counterexample. It is, however, possible to give constructions that extract multiple bits in a useful way, as we have shown in the previous section.

Which constructions give rise to "useful" extractors in a quantum context? In this section, we elaborate on this question, showing that general extractors can be used in the setting of privacy amplification *if* the adversary's memory is limited. Note that the setting of privacy amplification imposes less stringent requirements on the extractor than the setting of the bounded-storage model. Nevertheless, the only construction known to work in the quantum setting has been two-universal hashing [13]–[16].

While two-universal hashing has the advantage that it extracts all the randomness present in the source (i.e., the number of extracted bits can be as large as $H_\infty(X) - 2\log 1/\varepsilon$), it requires a long seed Y. Viewed as an extractor, two-universal hashing has the form $\mathsf{E} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$, i.e., the seed is of the same length as the source X. When applied to privacy amplification, this means that n bits need to be communicated from Alice to Bob. We will show below that by considering general extractors, the amount of communication can be reduced to the same order of magnitude as the number of qubits the adversary controls. This statement will be made more precise below.

We use a measurement-based approach which bounds the trace distance in terms of the outcomes of a tomographic measurement. More precisely, we use will use the following lemma, whose proof can be found in Appendix D.

Lemma 4: Let A be a Hermitian operator on $\mathbb{C}^d$. Then there is a POVM $\mathcal{F} = \{F_z\}_{z=1}^{d^2}$ such that

$$\|A\| \leq 2d^3 \cdot \|\mathcal{F}(A)\| \tag{38}$$

where $\mathcal{F}(A) = \sum_{z=1}^{d^2} \mathrm{tr}(F_z A) \cdot |z\rangle\langle z|$ for some orthonormal basis $\{|z\rangle\}_{z=1}^{d^2}$ of $\mathbb{C}^{d^2}$.

We point out that Lemma 4 is in general far from optimal and trivial for certain operators A. However, it has the advantage that the POVM $\mathcal{F}$ is independent of A. A few possible improvements are discussed in Appendix D.

We can use Lemma 4 to show the following.

Lemma 5: Let $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a strong (k,ε)-extractor. Then

$$d(\mathsf{E}(X,Y) \leftarrow YQ) \leq 4 \cdot 2^{3H_0(Q)} \cdot \varepsilon \tag{39}$$

for all cq-states ρ_{XQ} with

$$H_{\mathrm{g}}(X \leftarrow Q) \geq k + \log 1/\varepsilon.$$

Proof: By definition and (2)

$$d(\mathsf{E}(X,y) \leftarrow Q) = \sum_{z \in \mathcal{Z}} \left\| P_{\mathsf{E}(X,y)}(z)\rho_y^z - \frac{1}{|\mathcal{Z}|}\rho_Q \right\|$$

where ρ_y^z is the conditional state $\rho_y^z := \rho_{Q\,|\,\mathsf{E}(X,y)=z}$ for all $(y,z) \in \mathcal{Y} \times \mathcal{Z}$. By Lemma 4, we get

$$d(\mathsf{E}(X,y) \leftarrow Q) \leq 2^{3H_0(Q)+1} \sum_{z \in \mathcal{Z}} \left\| P_{\mathsf{E}(X,y)}(z)\mathcal{F}(\rho_y^z) - \frac{1}{|\mathcal{Z}|}\mathcal{F}(\rho_Q) \right\|$$

and thus by taking the expectation over $y \leftarrow P_Y$ with (2) (see also (44))

$$d(\mathsf{E}(X,Y) \leftarrow YQ) \leq 2^{3H_0(Q)+1}\, d(\mathsf{E}(X,Y) \leftarrow Y\mathcal{F}(Q)).$$

The claim then follows from Proposition $1'$. □

This lemma shows that in principle, any strong (k,ε)-extractor with suitable parameters can be used for privacy amplification. While the bound (39) will not give a nontrivial statement for parameters typical to the bounded-storage model (as it is based on Lemma 4), it allows us to reduce the amount of communication required in the setting of privacy amplification. We illustrate this using a construction by Srinivasan and Zuckerman [31] for simplicity, but we point out that using constructions from [32], it is possible to reduce the randomness required for privacy amplification even further. An efficiently computable strong (k,ε)-extractor $\mathsf{E} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ is constructed in [31] for any k, m, ε with $k \geq m + 2\log 1/\varepsilon + 2$, where $t = 2(k+m) + O(\log n)$. Applying this to a situation where the adversary is given at most $d \geq H_0(Q)$ qubits of storage, we obtain an efficiently computable function $\mathsf{E} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ which uses only

$$t = 4(3d + m + \log 1/\varepsilon + 3) + O(\log n) \tag{40}$$

bits of seed and satisfies $d(\mathsf{E}(X,Y) \leftarrow YQ) \leq \varepsilon$ whenever

$$H_\infty(X) \geq m + 10d + 3\log 1/\varepsilon + 8. \tag{41}$$

For certain parameters (i.e., sublinear bounds d on the adversary's storage), this construction is more efficient in terms of the seed length t than the local extractor described in Corollary 2. Note that, generally, we will have $d = \Omega(\log n)$, hence the number (40) of random bits used in this construction is dominated by the number of qubits the adversary controls, contrary to the two-universal hashing construction.

VII. Conclusion

While Holevo's celebrated theorem implies that n quantum bits cannot be used to store more than n classical bits reliably, this result is in general not applicable in cryptography, where even partial information can make a difference. Indeed, numerous examples are known where quantum bits are more powerful than the same number of classical bits (see e.g., [33], [34], [14], [20]). In this light, it is natural to study the potential advantage offered by quantum information with respect to specific tasks.

We have taken a step in this direction by showing that certain schemes for the bounded-storage model which are secure in the presence of classical adversaries are also secure in the presence of adversaries who are in control of quantum storage. Surprisingly, the corresponding security parameters are almost the same for the quantum and the classical case when only a single bit is extracted. It is straightforward to extend and reformulate this result in terms of communication complexity. It then states that there cannot be a large separation between the one-way average-case quantum and classical communication complexities of a Boolean function.

This is in sharp contrast to the case of extractors which output several bits. There are extractors that provide security in the classical bounded-storage model, but cannot safely be used against quantum adversaries [20]. Nevertheless, it is possible to give a family of constructions that yield secure bits; this is our main contribution.

While our extractors provide security against quantum adversaries, their parameters are far from optimal. Future work can focus on improving these constructions.

Appendix A
Properties of the Nonuniformity

We summarize a few properties of the nonuniformity in this section. The nonuniformity $d(Z \leftarrow W)$ of Z given W can be viewed as the average distance of the conditional distribution $P_{Z\,|\,W=w}$ to the uniform distribution, for a random choice of $w \leftarrow P_W$, that is

$$d(Z \leftarrow W) = \mathbb{E}_{w \leftarrow P_W}[d(Z\,|\,W=w)]. \tag{42}$$

More generally, for a ccq-state ρ_{ZWQ}, where W and Q are not necessarily independent, the nonuniformity of Z given WQ can be written as an average of the corresponding nonuniformities with respect to the conditional states $\rho_{ZQ\,|\,W=w}$. This is a direct consequence of (2). In formula, we have

$$d(Z \leftarrow WQ) = \mathbb{E}_{w \leftarrow P_W}[d(Z \leftarrow Q\,|\,W=w]. \tag{43}$$

In particular, we can write

$$d(\mathsf{E}(X,Y) \leftarrow YQ) = \mathbb{E}_{y \leftarrow P_{\mathcal{U}_\mathcal{Y}}}[d(\mathsf{E}(X,y) \leftarrow Q\,|\,Y=y] \tag{44}$$

where the term in brackets is equal to $d(\mathsf{E}(X,y) \leftarrow Q)$ when Y and Q are independent (which is usually the case in this paper).

Finally, we point out that conditioning on independent random variables leaves the nonuniformity invariant, that is

$$d(Z \leftarrow VQ) = d(Z \leftarrow Q) \tag{45}$$

if $\rho_{ZQV} = \rho_{ZQ} \otimes \rho_V$. This follows from (43).

APPENDIX B
PROOFS OF SECTION II

Proof: (**of Proposition 1**) Consider a random variable $\hat{X}$ defined by a channel $P_{\hat{X}\,|\,E}$ which takes a value $\hat{x}$ for which $P_{X\,|\,E=e}(\hat{x}) = 2^{-H_\infty(X\,|\,E=e)}$ with certainty, for every $e \in \mathcal{E}$. Clearly, we have (cf. (6))

$$2^{-H_{\mathsf{g}}(X\leftarrow E)} = \Pr[X = \hat{X}] = \mathbb{E}_{e\leftarrow P_E}\left[2^{-H_\infty(X\,|\,E=e)}\right].$$

By Markov's inequality, this implies that

$$\Pr_{e\leftarrow P_E}[H_\infty(X\,|\,E=e) \le H_{\mathsf{g}}(X\leftarrow E) - \log 1/\varepsilon] \le \epsilon.$$

The result then follows by convexity, using (4) and the fact that

$$d(\mathsf{E}(X,Y)\leftarrow YE) = \mathbb{E}_{e\leftarrow P_E}[d(\mathsf{E}(X,Y)\leftarrow Y\,|\,E=e]$$

because of (42). □

Lemma 1 can be seen as a special case of Lemma 1′. Their proofs are analogous, but we include both here, as the classical proof is instructive for the quantum generalization.

Proof: (**of Lemma 1**) Let $a_{v,w} := 2^{-H_{\mathsf{g}}(X\leftarrow E\,|\,V=v,W=w)}$ for all $(v,w) \in \mathcal{V}\times\mathcal{W}$. By definition

$$a_{v,w} = \sum_{e\in\mathcal{E}} P_{E\,|\,V=v,W=w}(e)\max_{x\in\mathcal{X}} P_{X\,|\,E=e,V=v,W=w}(x).$$

In particular

$$\begin{aligned} &P_{VW}(v,w)a_{v,w} \\ &\quad= \sum_{e\in\mathcal{E}} P_{EVW}(e,v,w)\max_{x\in\mathcal{X}} P_{X\,|\,E=e,V=v,W=w}(x) \\ &\quad= \sum_{e\in\mathcal{E}} \max_{x\in\mathcal{X}} P_{XEVW}(x,e,v,w). \end{aligned}$$

But by summing over $w \in \mathcal{W}$

$$P_{XEVW}(x,e,v,w) \le P_{XEV}(x,e,v)$$

and thus

$$\begin{aligned} &P_{VW}(v,w)a_{v,w} \\ &\quad\le P_V(v)\sum_{e\in\mathcal{E}} P_{E\,|\,V=v}(e)\max_{x\in\mathcal{X}} P_{X\,|\,V=v,E=e}(x) \\ &\quad= P_V(v)\sum_{e\in\mathcal{E}} P_E(e)\max_{x\in\mathcal{X}} P_{X\,|\,E=e}(x) \\ &\quad= P_V(v)2^{-H_{\mathsf{g}}(X\leftarrow E)} \end{aligned}$$

where we used the independence of XE and V in the second step and the definition of $H_{\mathsf{g}}(X\leftarrow E)$ to obtain the last identity. We conclude that

$$\mathbb{E}_{(v,w)\leftarrow P_{VW}}[a_{v,w}] \le 2^{H_0(W)-H_{\mathsf{g}}(X\leftarrow E)}.$$

It is easy to see that

$$\mathbb{E}_{(v,w)\leftarrow P_{VW}}[a_{v,w}] = 2^{-H_{\mathsf{g}}(X\leftarrow VWE)}$$

which proves our first claim. We then use Markov's inequality to obtain

$$\Pr_{(v,w)\leftarrow P_{VW}}\left[a_{v,w} \ge \frac{1}{\varepsilon}2^{H_0(W)-H_{\mathsf{g}}(X\leftarrow E)}\right] \le \varepsilon$$

which is our second claim.

Proof: (**of Lemma 1′**) By assumption, ρ_{XVWQ} has the form

$$\rho_{XVWQ} = \sum_{(x,v,w)\in\mathcal{X}\times\mathcal{V}\times\mathcal{W}} P_{XVW}(xvw)|xvw\rangle\langle xvw|\otimes\rho_x.$$

For every $(v,w) \in \mathcal{V}\times\mathcal{W}$, let $\mathcal{E}^{v,w} := \{E_x^{v,w}\}_{x\in\mathcal{X}}$ be the POVM which maximizes the expression in the definition of $H_{\mathsf{g}}(X \leftarrow Q\,|\,V=v,W=w)$. We define the operators $\{F_x^v\}_{x\in\mathcal{X}}$ by

$$F_x^v := 2^{-H_0(W)}\sum_{w\in\mathcal{W}} E_x^{v,w}.$$

It is easy to see that $\mathcal{F}^v := \{F_x^v\}_{x\in\mathcal{X}}$ forms a POVM for every $v \in \mathcal{V}$, and the operator inequality $E_x^{v,w} \le 2^{H_0(W)}F_x^v$ holds. In particular

$$\mathsf{tr}(E_x^{v,w}\rho_x) \le 2^{H_0(W)}\mathsf{tr}(F_x^v\rho_x) \tag{46}$$

for all $(x,v,w) \in \mathcal{X}\times\mathcal{V}\times\mathcal{W}$. Let us introduce the abbreviation

$$a_{v,w} := 2^{-H_{\mathsf{g}}(X\leftarrow Q\,|\,V=v,W=w)} \tag{47}$$

for every $(v,w) \in \mathcal{V}\times\mathcal{W}$. By definition and (46)

$$\begin{aligned} a_{v,w} &= \sum_{x\in\mathcal{X}} P_{X\,|\,V=v,W=w}(x)\mathsf{tr}(E_x^{v,w}\rho_x) \\ &\le 2^{H_0(W)}\sum_{x\in\mathcal{X}} P_{X\,|\,V=v,W=w}(x)\mathsf{tr}(F_x^v\rho_x). \end{aligned}$$

We thus have

$$\begin{aligned} &\mathbb{E}_{(v,w)\leftarrow P_{VW}}[a_{v,w}] \\ &\quad\le 2^{H_0(W)}\sum_{(x,v)\in\mathcal{X}\times\mathcal{V}} P_{XV}(x,v)\mathsf{tr}\left(F_x^v\rho_x\right) \\ &\quad= 2^{H_0(W)}\mathbb{E}_{v\leftarrow P_V}\left[\sum_{x\in\mathcal{X}} P_{X\,|\,V=v}(x)\mathsf{tr}(F_x^v\rho_x)\right]. \end{aligned} \tag{48}$$

But for every $v \in \mathcal{V}$

$$\begin{aligned} \sum_{x\in\mathcal{X}} P_{X\,|\,V=v}(x)\mathsf{tr}\left(F_x^v\rho_x\right) &= \sum_{x\in\mathcal{X}} P_X(x)\mathsf{tr}\left(F_x^v\rho_x\right) \\ &\le 2^{-H_{\mathsf{g}}(X\leftarrow Q)} \end{aligned} \tag{49}$$

by the assumption that $\rho_{XV} = \rho_X\otimes\rho_V$, and the definition of the latter quantity. Combining (49) with (48) and (47) gives

$$2^{-H_{\mathsf{g}}(X\leftarrow VWQ)} = \mathbb{E}_{(v,w)\leftarrow P_{VW}}[a_{v,w}] \le 2^{H_0(W)-H_{\mathsf{g}}(X\leftarrow Q)}$$

and the first claim follows. The second claim follows from Markov's inequality, as in the Proof of Lemma 1. □

APPENDIX C
PROOF OF REFINEMENT LEMMA

In the proof of Theorem 1, we have used the fact that applying classical post-processing after a measurement $\mathcal{F}$ does not increase the nonuniformity. We state this as a lemma; the proof is trivial and follows from the triangle inequality.

Lemma 6: Let $P_{E\,|\,X}$ be a channel, and let $\mathcal{F} := \{F_x\}_{x\in\mathcal{X}}$ be a POVM on $\mathcal{Q}$. Define the operators

$$E_e := \sum_{x\in\mathcal{X}} P_{E\,|\,X=x}(e)F_x$$

for every $e \in \mathcal{E}$. Then $\mathcal{E} := \{E_e\}_{e\in\mathcal{E}}$ is a POVM, and for any cq-state ρ_{ZQ},

$$d(Z \leftarrow \mathcal{E}(Q)) \leq d(Z \leftarrow \mathcal{F}(Q)).$$

Proof: It is trivial to check that $\mathcal{E}$ is indeed a POVM. By definition

$$\begin{aligned} d(Z \leftarrow \mathcal{E}(Q)) &= \left\|\rho_{Z\mathcal{E}(Q)} - \rho_{\mathcal{U}_\mathcal{Z}} \otimes \rho_{\mathcal{E}(Q)}\right\| \\ &= \frac{1}{2}\sum_{(z,e)\in\mathcal{Z}\times\mathcal{E}} \alpha_{z,e} \end{aligned} \tag{50}$$

where

$$\alpha_{z,e} := \|\mathsf{tr}((|z\rangle\langle z| \otimes E_e)\rho_{ZQ}) - \frac{1}{|\mathcal{Z}|}\mathsf{tr}(E_e\rho_Q)\|.$$

By the definition of E_e and the triangle inequality

$$\alpha_{z,e} \leq \sum_{x\in\mathcal{X}} P_{E|X=x}(e)\left\|\mathsf{tr}((|z\rangle\langle z| \otimes F_x)\rho_{ZQ}) - \frac{1}{|\mathcal{Z}|}\mathsf{tr}(F_x\rho_Q)\right\|.$$

Combining this with (50) gives the claim. □

Appendix D
Informationally Complete POVMs and the Trace Norm

In this appendix, we give a proof of Lemma 4 and discuss possible improvements. We will use an informationally complete POVM, i.e., a family of operators which is both a POVM and a basis of the set of Hermitian operators on $\mathbb{C}^d$. Note that the latter set forms a real d^2-dimensional Hilbert space with inner product $(A, B) = \mathsf{tr}(AB)$. In particular, there is a well-defined notion of a dual space. We will use the following terminology. A family of Hermitian operators $\{F_z\}_{z\in\mathcal{Z}}$ is dual to a family of Hermitian operators $\{F_z^*\}_{z\in\mathcal{Z}}$ if

$$A = \sum_{z\in\mathcal{Z}} \mathsf{tr}(F_z A)F_z^* \tag{51}$$

for all Hermitian operators A. This identity implies that every operator is completely specified by the values $\mathsf{tr}(F_z A)$, $z \in \mathcal{Z}$, and these depend linearly on A. In [35, Lemmas III.5–III.8], pairs of families $\{F_z^*\}_{z=1}^{d^2}$ and $\mathcal{F} = \{F_z\}_{z=1}^{d^2}$ satisfying (51) are constructed for any dimension d with the additional property that $\mathcal{F}$ is a POVM and

$$\mathsf{tr}(|F_z^*|) \leq \sqrt{2}d^3, \quad \text{for all } z = 1, \ldots, d^2. \tag{52}$$

Lemma 4 now directly follows from (51), (52), and the triangle inequality (we bound $\sqrt{2}$ by 2 for convenience).

The statement of Lemma 4 can be improved in several ways under additional assumptions. For example, if so-called symmetric informationally complete POVMs exist in $\mathbb{C}^d$, then the exponent in (38) can be improved from 3 to 1 (cf. [35, Lemma 3.2]). We have recently learned [36] that this improvement can be obtained for general dimension d using the t-design constructions by Ambainis and Emerson [37]. In our application, this reduces the required amount of seed (40) and min-entropy (41), but does not affect the qualitative nature of our results.

Here we would like to point out that Lemma 4 can be generalized to a situation where part of the operator A is "classical." The corresponding statement then has the following form.

Lemma 4′: Let $\{|i\rangle\}_{i=1}^{d_1}$ be an orthonormal basis of $\mathbb{C}^{d_1}$, and let A be a Hermitian operator on $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ of the form $A = \sum_i |i\rangle\langle i| \otimes A_i$. Then there is a POVM $\mathcal{F} = \{F_{(i,z)}\}_{(i,z)\in\{1,\ldots,d_1\}\times\{1,\ldots,d_2^2\}}$ such that

$$\|A\| \leq 2d_2^3 \cdot \|\mathcal{F}(A)\| \tag{53}$$

where

$$\mathcal{F}(A) = \sum_{i=1}^{d_1}\sum_{z=1}^{d_2^2} \mathsf{tr}(F_{(i,z)}A) \cdot |i\rangle\langle i| \otimes |z\rangle\langle z|$$

for some orthonormal basis $\{|z\rangle\}_{z=1}^{d_2^2}$ of $\mathbb{C}^{d_2^2}$.

This variation of Lemma 4 is obtained simply by setting $F_{(i,z)} := |i\rangle\langle i| \otimes F_z$, where F_z is the POVM used previously. It is then easy to see to (51) holds for all operators A of the specified form with F_z^* replaced by $F_{(i,z)}^* = |i\rangle\langle i| \otimes F_z^*$. The claim follows since $\mathsf{tr}(|F_{(i,z)}^*|) = \mathsf{tr}(|F_z^*|)$.

Lemma 4′ implies that giving the adversary classical information E in addition to Q does not increase the distance $d(\mathsf{E}(X, Y) \leftarrow YEQ)$ by more than what is implied by the amount of prior information $H_\mathsf{g}(X \leftarrow EQ)$. In particular, we obtain the following generalization of Lemma 5.

Lemma 5′: Let $\mathsf{E} : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a strong (k, ε)-extractor. Then

$$d(\mathsf{E}(X, Y) \leftarrow YEQ) \leq 4 \cdot 2^{3H_0(Q)} \cdot \varepsilon \tag{54}$$

for all ccq-states ρ_{XEQ} with

$$H_\mathsf{g}(X \leftarrow EQ) \geq k + \log 1/\varepsilon.$$

As expected, classical information "behaves classically" with respect to extractors. We do not elaborate on this further, as this is not the main focus of our work. We emphasize, however, that the case of hybrid classical-quantum prior information is implicitly already covered by our results. In concrete situations, the analysis boils down to obtaining estimates on the quantity $H_\mathsf{g}(X \leftarrow Q)$, where Q may consist of both classical and quantum parts.

Acknowledgment

R. T. König wishes to thank Ueli Maurer and Renato Renner for interesting discussions about bounded-storage cryptography. He would also like to thank IBM T. J. Watson Research Center for their hospitality during his stay there. B. M. Terhal would like to thank Yevgeniy Dodis and Roberto Oliveira for many discussions on the security of the bounded-storage model. The authors wish to thank Ronald de Wolf for helpful comments, in particular in relation to Remark 1. They also thank Yevgeniy Dodis for the suggestion to consider independent

randomizers, and the reviewers for their detailed and helpful comments.

REFERENCES

[1] N. Nisan and D. Zuckerman, "Randomness is linear in space," *J. Comp. Syst. Sci.*, vol. 52, pp. 43–52, 1996, a preliminary version was presented at STOC'93.

[2] R. Shaltiel, "Recent developments in explicit constructions of extractors," *Bull. EATCS* vol. 77, pp. 67–95, 2002 [Online]. Available: http://dblp.uni-trier.de

[3] G. S. Vernam, "Cipher printing telegraph systems for secret wire and radio telegraphic communications," *J. Amer. Inst. Elec. Eng.*, vol. 55, pp. 109–115, 1926.

[4] B. Pfitzmann and M. Waidner, "Composition and integrity preservation of secure reactive systems," in *Proc. 7th ACM Conf. Computer and Communications Security*, 2000, pp. 245–254.

[5] R. Canetti, "Security and composition of multi-party cryptographic protocols," *J. Cryptol.*, vol. 13, no. 1, pp. 143–202, 2000.

[6] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, no. 2, pp. 210–229, 1988.

[7] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pt. 2, pp. 1915–1923, Nov. 1995.

[8] U. Maurer, "Conditionally-perfect secrecy and a provably-secure randomized cipher," *J. Cryptol.*, vol. 5, no. 1, pp. 53–66, 1992.

[9] Y. Aumann, Y. Z. Ding, and M. O. Rabin, "Everlasting security in the bounded storage model," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1668–1680, Jun. 2002.

[10] S. Dziembowski and U. Maurer, "Tight security proofs for the bounded-storage model," in *Proc. 34th Annu. ACM Symp. Theory of Computing*, 2002, pp. 341–350, ser. Lecture Notes in Computer Science.

[11] C. Lu, "Hyper-encryption against space-bounded adversaries from on-line strong extractors," in *Advances in Cryptology—CRYPTO 2003*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-verlag, 2002, pp. 18–22.

[12] S. P. Vadhan, "On constructing locally computable extractors and cryptosystems in the bounded-storage model," *J. Cryptology*, vol. 17, no. 1, pp. 43–77, 2004.

[13] M. Ben-Or, "Security of BB84 QKD Protocol," 2002 [Online]. Available: http://www.msri.org/publications/ln/msri/2002/quantumintro, (slides) Part II.

[14] R. König, U. Maurer, and R. Renner, "On the power of quantum memory," *IEEE Trans. Inf. Theory* vol. 51, no. 7, pp. 2391–2401, Jul. 2005 [Online]. Available: http://arxiv.org/abs/quant-ph/0305154

[15] R. Renner and R. König, "Universally composable privacy amplification against quantum adversaries," in *Proc. Second Theory of Cryptography Conf. TCC*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2005, vol. 3378, pp. 407–425 [Online]. Available: http://arxiv.org/abs/quant-ph/0403133

[16] R. Renner, "Security of quantum key distribution" Ph.D. dissertation, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 2005 [Online]. Available: http://arxiv.org/abs/quant-ph/0512258

[17] M. Christandl, R. Renner, and A. Ekert, A Generic Security Proof for Quantum Key Distribution Feb. 2004 [Online]. Available: http://arxiv.org/abs/quant-ph/0402131

[18] I. Damgård, S. Fehr, L. Salvail, and C. Schaffner, "Cryptography in the bounded quantum-storage model," in *Proc. 46th Annu. IEEE Symp. Foundations of Computer Science (FOCS)*, Pittsburgh, PA, Oct. 2005, pp. 449–458.

[19] H. Buhrman, M. Christandl, P. Hayden, H. W. Lo, and S. Wehner, "On the (Im)Possibility of Quantum String Commitment," 2005 [Online]. Available: http://arxiv.org/abs/quant-ph/0504078

[20] D. Gavinsky, J. Kempe, and R. de Wolf, "Exponential Separation of Quantum and Classical One-Way Communication Complexity for a Boolean Function," 2006 [Online]. Available: http://arxiv.org/abs/quant-ph/0607174

[21] S. Wolf, "Information-Theoretically and Computationally Secure Key Agreement in Cryptography" Ph.D. dissertation, Swiss Federal Institute of Technology (ETH Zurich), Zurich, Switzerland, 1999, ETH dissertation No. 13138.

[22] D. DiVincenzo, M. Horodecki, D. Leung, J. Smolin, and B. Terhal, "Locking classical correlations in quantum states," *Phys. Rev. Lett.*, vol. 92, p. 067902, 2004.

[23] R. König, R. Renner, A. Bariska, and U. Maurer, "Small accessible quantum information does not imply security," *Phys. Rev. Lett.* vol. 98, no. 14, p. 140502, 2007 [Online]. Available: http://link.aps.org/abstract/PRL/v98/e140502

[24] D. Unruh, "Formal Security in Quantum Cryptology" Master's thesis, Karlsruhe, Germany, Dec. 2002 [Online]. Available: http://www.unruh.de/DniQ/publications

[25] D. Unruh, Simulatable Security for Quantum Protocols Sept. 2004 [Online]. Available: http://arxiv.org/abs/quant-ph/0409125

[26] M. Ben-Or and D. Mayers, General Security Definition and Composability for Quantum and Classical Protocols Sep. 2004 [Online]. Available: http://arxiv.org/abs/quant-ph/0409062

[27] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, "The universal composable security of quantum key distribution," in *Proc. Second Theory of Cryptography Conf. TCC*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2005, vol. 3378, pp. 386–406 [Online]. Available: http://arxiv.org/abs/quant-ph/0409078

[28] H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," *J. Math. Phys.*, vol. 43, no. 5, pp. 2097–2106, 2002.

[29] P. Hausladen and W. Wootters, "A 'pretty good' measurement for distinguishing quantum states," *J. Mod. Opt.*, vol. 41, no. 12, pp. 2385–2390, 1994.

[30] C. W. Helstrom, *Quantum Detection and Estimation Theory*. New York: Academic, 1976.

[31] A. Srinivasan and D. Zuckerman, "Computing with very weak random sources," in *Proc. IEEE Symp. Foundations of Computer Science*, Santa Fe, NM, Nov. 1994, pp. 264–275 [Online]. Available: citeseer.ist.psu.edu/447609.html

[32] O. Reingold, R. Shaltiel, and A. Wigderson, "Extracting randomness via repeated condensing," in *Proc. IEEE Symp Foundations of Computer Science*, Redondo Beach, CA, Nov. 2000, pp. 22–31 [Online]. Available: citeseer.ist.psu.edu/reingold00extracting.html

[33] A. Nayak, "Optimal lower bounds for quantum automata and random access codes," in *Proc. 40th Annu. Symp. Foundations of Computer Science*, New York City, Oct. 1999, pp. 369–377, quant-ph/9904093.

[34] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, "Dense quantum coding and a lower bound for 1-way quantum automata," in *Proc. 31th ACM Symp. Theory of Computing*, Atlanta, GA, May 1999, quant-ph/9804043.

[35] R. König and R. Renner, "A de Finetti representation for finite symmetric quantum states," *J. Math. Phys.*, vol. 46, p. 122108, 2005.

[36] P. Sen and A. Nayak, 2007, personal communication.

[37] A. Ambainis and J. Emerson, Quantum t-Designs: T-Wise Independence in the Quantum World 2007 [Online]. Available: http://arxiv.org/abs/quant-ph/0701126